

**PCT**WORLD INTELLECTUAL PROPERTY ORGANIZATION  
International Bureau

## INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 6: <b>C12N 15/11, C12Q 1/68, C07K 14/40, 16/14, G01N 33/569, A61K 39/00</b>	A1	(11) International Publication Number: <b>WO 96/36707</b> (43) International Publication Date: 21 November 1996 (21.11.96)
(21) International Application Number: <b>PCT/IT96/00097</b>		(74) Agents: DE SIMONE, Domenico et al.; Ing. Barzano' & Zanardo Roma S.p.A., Via Piemonte, 26, I-00187 Roma (IT).
(22) International Filing Date: 15 May 1996 (15.05.96)		(81) Designated States: AL, AM, AT, AU, AZ, BB, BG, BR, BY, CA, CH, CN, CZ, DE, DK, EE, ES, FI, GB, GE, HU, IS, JP, KE, KG, KP, KR, KZ, LK, LR, LS, LT, LU, LV, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, TJ, TM, TR, TT, UA, UG, US, UZ, VN, ARIPO patent (KE, LS, MW, SD, SZ, UG), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, ML, MR, NE, SN, TD, TG).
(30) Priority Data: <b>RM95A000314 16 May 1995 (16.05.95) IT</b>		(71) Applicants (for all designated States except US): UNIVERSITA' DEGLI STUDI DI ROMA "LA SAPIENZA" [IT/IT]; Piazzale Aldo Moro, 5, I-00185 Roma (IT). ISTITUTO SUPERIORE DI SANITA' [IT/IT]; Viale Regina Elena, 299, I-00161 Roma (IT).
(72) Inventors; and (75) Inventors/Applicants (for US only): CASSONE, Antonio [IT/IT]; Istituto Superiore di Sanita', Viale Regina Elena, 299, I-00161 Roma (IT). LA VALLE, Roberto [IT/IT]; Istituto Superiore di Sanita', Viale Regina Elena, 299, I-00161 Roma (IT). BROMURO, Carla [IT/IT]; Istituto Superiore di Sanita', Viale Regina Elena, 299, I-00161 Roma (IT). CRISANTI, Andrea [IT/IT]; Universita' degli Studi di Roma "La Sapienza", Piazzale Aldo Moro, 5, I-00185 Roma (IT). MULLER, Hans, Michael [IT/IT]; Universita' degli Studi di Roma "La Sapienza", Piazzale Aldo Moro, 5, I-00185 Roma (IT).		(73) Published With international search report.

(54) Title: CANDIDA HEAT SHOCK PROTEIN, cDNA AND USES THEREOF

## (57) Abstract

A nucleotide sequence and related protein from Candida homologous to 70 kd heat shock protein, for uses in diagnosis and therapy.

***FOR THE PURPOSES OF INFORMATION ONLY***

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AM	Armenia	GB	United Kingdom	MW	Malawi
AT	Austria	GE	Georgia	MX	Mexico
AU	Australia	GN	Guinea	NE	Niger
BB	Barbados	GR	Greece	NL	Netherlands
BE	Belgium	HU	Hungary	NO	Norway
BF	Burkina Faso	IE	Ireland	NZ	New Zealand
BG	Bulgaria	IT	Italy	PL	Poland
BJ	Benin	JP	Japan	PT	Portugal
BR	Brazil	KE	Kenya	RO	Romania
BY	Belarus	KG	Kyrgyzstan	RU	Russian Federation
CA	Canada	KP	Democratic People's Republic of Korea	SD	Sudan
CF	Central African Republic	KR	Republic of Korea	SE	Sweden
CG	Congo	KZ	Kazakhstan	SG	Singapore
CH	Switzerland	LI	Liechtenstein	SI	Slovenia
CI	Côte d'Ivoire	LK	Sri Lanka	SK	Slovakia
CM	Cameroon	LR	Liberia	SN	Senegal
CN	China	LT	Lithuania	SZ	Swaziland
CS	Czechoslovakia	LU	Luxembourg	TD	Chad
CZ	Czech Republic	LV	Larvia	TG	Togo
DE	Germany	MC	Monaco	TJ	Tajikistan
DK	Denmark	MD	Republic of Moldova	TT	Trinidad and Tobago
EE	Estonia	MG	Madagascar	UA	Ukraine
ES	Spain	ML	Mali	UG	Uganda
FI	Finland	MN	Mongolia	US	United States of America
FR	France	MR	Mauritania	UZ	Uzbekistan
GA	Gabon			VN	Viet Nam

CANDIDA HEAT SHOCK PROTEIN, cDNA AND USES THEREOF

The invention concerns the cDNA and the corresponding protein of a heat shock protein isolated from C. albicans, and fragments thereof to develop methods to identify C. albicans in biological and/or environment samples, and/or preparations either for therapeutic, prophylaxis or vaccine purpose.

Pathogenic yeasts are the major agents of opportunistic infections in immunosuppressed patients, in particular AIDS, tumor, neutropenia patients or bone marrow transplanted subjects (1). HIV<sup>+</sup> subject susceptibility to C. albicans is related to the strong decrease of cell-mediated immunity because of the numerical and functional decrease of CD4<sup>+</sup> helper-inducer lymphocytes (2).

C. albicans cell wall mannoproteins and heatshock proteins of other microorganisms as well, are major antigens and immunomodulators, and play a relevant role during host invasion and infection (3,4).

By using a rabbit immune serum obtained against heat-inactivated C. albicans ATCC 20955 strain cells, the authors of the instant invention isolated the caRLV130 clone from an expression library in the λgt11 phage obtained by cDNA isolated from C. albicans at the yeast growth stage. Said clone contains a DNA insert of 2325 base pairs which codes in the 5'-3' direction from +105 to +2072 for a 656 aminoacid protein having a strong homology with a S. cerevisiae heat shock protein 70.

HSPs are induced by different stresses, either chemical or physical, normally by heating. Many HSPs are present and active also in non stressed cells, where they play important functions of cell physiology ("chaperonins"). They may be grouped in families of

different molecular weights, very conserved even among phylogenesis distant organisms (5). Therefore it should not be surprising either that HSPs are involved in the immune response, or that they represent major antigens of different pathogenic agents, or that they may give autoimmune responses, given to the fact that the infection itself represents an extreme form of stress, both for the infectious agent and for the host (4).

It is therefore an object of the invention a nucleic acid comprising a nucleotide sequence coding the protein having the amino acid sequence of SEQ ID No.2 or parts thereof. Preferably the nucleic acid comprises a nucleotide sequence with at least a 65% homology with the nucleotide sequence of SEQ ID No.1 or parts thereof. More preferably the nucleic acid comprises the nucleotide sequence of SEQ ID No.1 or parts thereof.

Further object of the invention is a composition comprising a nucleic acid comprising a nucleotide sequence coding the protein having the amino acid sequence of SEQ ID No.1 or parts thereof. Preferably the composition comprises a nucleic acid having a nucleotide sequence with at least a 65% homology with the nucleotide sequence of SEQ ID No.1 or parts thereof. More preferably the composition comprises a nucleic acid having the nucleotide sequence of SEQ ID No.1 or parts thereof.

Further object of the invention is the use of the nucleic acid comprising a nucleotide sequence coding the protein having the amino acid sequence of SEQ ID No.2 or parts thereof for oligonucleotide probes to be used in diagnosis and typing of Candida related pathologies. The use of a nucleic acid having at least a 65% homology with the nucleotide sequence of SEQ ID No.1 or parts thereof is preferred. The use of a nucleic acid having the

nucleotide sequence of SEQ ID No.1 or parts thereof is most preferred.

The oligonucleotides of the invention are advantageously used for PCR (polymerase chain reaction) 5 to detect the presence in biological and/or environment samples either of C. albicans or of other Candida species or of yeast-like related microorganisms comprising said gene; in a labeled form (radionuclides, biotin, enzymes, etc.) to detect the presence in biological and/or 10 environment samples either of C. albicans or of other related; for the C. albicans or related species typing and/or diagnosis; as potential antibiotic and/or chemiotherapeutic targets, or antisense RNA active for Candida species and/or yeast-like related microorganisms 15 coding an homologous sequence.

Another object of the invention is a polypeptide having the aminoacid sequence comprised in the SEQ ID No.2, or having at least a 50% homology with SEQ ID No. 2 or fragments, and/or functional and immunologic 20 homologous thereof.

Further object of the invention is a composition comprising a polypeptide having an amino acid sequence comprised in SEQ ID No.2 or having at least a .50% homology with SEQ ID No. 2 or fragments, and/or 25 functional and immunologic homologous thereof.

Further object of the invention is the use of a polypeptide having the amino acid sequence comprised in SEQ ID No.2 or having at least a 50% homology with SEQ ID No. 2 or of fragments, and/or functional and/or 30 immunologic homologous thereof to make polyclonal or monoclonal antibodies against the 70 kd heat shock protein (HSP70) of C. albicans or related species.

Further object of the invention is the use of a polypeptide having the amino acid sequence comprised in

SEQ ID No.2 or having at least a 50% homology with SEQ ID No. 2 or of fragments, and/or functional and/or immunologic homologous thereof to detect C. albicans and related species HSP70 in a biological sample having a 5 human, animal or environmental origin.

Further object of the invention is the use of a polypeptide having the amino acid sequence comprised in SEQ ID No.2 or having at least a 50% homology with SEQ ID No. 2 or of fragments, and/or functional and/or 10 immunologic homologous thereof for the preparation of a composition to be used for prophylaxis and/or therapy of C. albicans or related microorganisms (pathogenic yeasts) diseases.

Further object of the invention is the use of a 15 polypeptide having the amino acid sequence comprised in SEQ ID No.2 or having at least a 50% homology with SEQ ID No. 2 or of fragments, and/or as potential antibiotic and/or chemiotherapeutic targets active for Candida species and/or yeast-like related microorganisms coding an 20 homologous sequence.

The invention will be described in different embodiments for clarifying but not limiting purposes.

Figure 1 represents the 1971 base pair DNA sequence (small letters) corresponding to the open reading frame 25 of λgt11-(caRLA130) clone insert and deduced aminoacid sequence (capital letters one-letter code).

Figure 2 represents the nucleotide sequence of the coding insert of caRLV130 clone (small letter) and comparison with S. cerevisiae YSCSSA1 gene (capital 30 letter).

Figure 3 represents the 656 aminoacid sequence deduced from the coding insert of caRLV130 clone (small letter) and comparison with the S. cerevisiae YSCSSA1

gene (capital letter). The aminoacid code utilized is the one letter code.

Figure 4 represents in panel A. Southern blot analysis of C. albicans strain ATCC 20955 chromosomes, obtained by pulse field electrophoresis (TAFE). The caRLV130 probe labeling refers to the highest molecular weight chromosome (3.5 Mbp). In panel B. Electrophoretic separation of C. albicans strain ATCC 20955 chromosomes.

Figure 5 represents on the left side: Northern blot analysis by hybridization of total RNA extracted from C. albicans cells grown at 22°C and transferred at 37°C for the time indicated with radiolabeled caRLV130 (cahsp70) and actin probes. The actin probe hybridization was performed to control the RNA amount on filters (see ref. 8). On the right side: immunoblotting reactivity of anti-CAHSP70 mouse serum with C. albicans extracts, at different times further to inducing a heat shock response as previously described.

Figure 6 represents in panel A. SDS-PAGE analysis:  
20 a) expression products of E. coli M15 containing the pDS56/RBS-E'-6his caRLV130/1 plasmid; b) expression products of E. coli M15 containing the pDS56/RBS-E'-6his caRLV130/2 plasmid; c) expression products of E. coli M15 containing the pDS56/RBS-E'-6his caRLV130/3 plasmid; d) expression products of E. coli M15 containing the pDS56/RBS-E'-6his caRLV130/4 plasmid. N.I.: Non induced E. coli culture extracts. I.: 1 mM IPTG induced E. coli culture extracts. P.: Purified fraction on histidine affinity nickel column from 1 mM IPTG induced E. coli culture extracts. In panel B. Schematic representation of caRLV130 coding sequence portions cloned into recombinant plasmids used in panel A. Right side: molecular weight in kDa. Left side: denomination of the expression product of recombinant plasmid. For further details, see table I.

Figure 7 represents the reactivity after immunoblotting on nitrocellulose filters of mouse sera as shown in the figure obtained against CAHSP70 fragments; a) expression products of nickel column purified pDS56/RBSII-E'-6his caRLV130/1 plasmid in 1 mM IPTG induced E. coli; b) expression products of nickel column purified pDS56/RBSII-E'-6his caRLV130/2 plasmid in 1 mM IPTG induced M15 E. coli; c) expression products of nickel column purified pDS56/RBSII-E'-6his caRLV130/3 plasmid in 1 mM IPTG induced M15 E. coli; d) expression products of nickel column purified pDS56/RBSII-E'-6his caRLV130/4 plasmid in 1 mM IPTG induced M15 E. coli (see also Fig. 6 and table I for a definition of polypeptide fragments). Left side: molecular weight of purified fragments.

Figure 8 represents the reactivity after immunoblotting on nitrocellulose filters of wealthy human sera obtained against CAHSP70 and fragments thereof; a) expression products of nickel column purified pDS56/RBSII-E'-6his caRLV130/1 plasmid in 1 mM IPTG induced M15 E. coli; b) expression products of nickel column purified pDS56/RBSII-E'-6his caRLV130/2 plasmid in 1 mM IPTG induced M15 E. coli; c) expression products of nickel column purified pDS56/RBSII-E'-6his caRLV130/3 plasmid in 1 mM IPTG induced M15 E. coli; d) expression products of nickel column purified pDS56/RBSII-E'-6his caRLV130/4 plasmid in 1 mM IPTG induced M15 E. coli. Left side: molecular weight of purified fragments. Right side: denomination of purified protein fragments. For further details see also table I.

Figure 9 represents in panel A. PCR experiment performed using oligonucleotide combination CA2-CA3 in the presence of C. albicans, C. parapsilosis (2), C. glabrata (3), C. guillermondii (4), C. krusei (5), C.

tropicalis (6), Mus muris (7), E. coli (8), S. cerevisiae (9) DNAs. Control with no DNA is as (10). At the right side the molecular weight of the amplified fragment is indicated. In panel B. PCR experiment using the 5 combination of CA1-CA4 oligonucleotides in the presence of C. albicans cDNA: DNA amplified from C. albicans DNA: 10 ng (2); 1 ng (3); 100 pg (4); 10 pg (5); 1 pg (6). Control: reaction with no DNA (1). PCR reaction conditions are as follows: 90 sec. 94°C denaturation; 90 10 sec. 60°C annealing; 120 sec. 72°C extension; 25 cycles.

Figure 1 shows the 1971 bp coding region of the isolated gene.

The caRLV130 sequence was filed with EMBL data base (No. Z30210). No intron can be found in the intronic 15 sequence, as shown by PCR product analysis and by "Southern-blot". By comparing the caRLV130 insert sequence with sequences present in the 6.7 version "GENE BANK" data base, some homologies can be detected. The insert shows the most high homology with the S. 20 cerevisiae gene SSA1 (one of the nine heatshock yeast gene family). The overall nucleotide sequence homology is of 78.8% in the coding region (figg. 2 and 3).

The gene corresponding to the caRLV130 sequence was mapped on the C. albicans chromosome showing the highest 25 molecular weight (3.5 Mpb) by pulse field electrophoresis (transverse-alternate: TAFE) utilizing the caRLV130 labeled cDNA insert as hybridization probe with C. albicans chromosomes blotted on nitrocellulose filters (fig. 4A an 4B). Gene transcription is activated by 30 exposing cells to a temperature higher than room temperature (thermal shift from 22°C to 37°C). Such finding was demonstrated by hybridization experiments using C. albicans total RNA (from cells grown either at 22°C or at 37°C, fractionated according to molecular

weight on formaldehyde agarose gel and blotted on nitrocellulose filters) and the caRLV130 DNA insert as radioactive probe. The induction of transcription is coupled also to an increase of protein expression, 2, 6 5 and 24 hours further to the 22°C to 37°C temperature shift (see fig. 5).

Different portions of the caRLV130 insert sequence were cloned in the expression plasmid pDS56/RBSII-E-6his (6), and coded polypeptides were expressed in E. coli 10 after fusion of their amino terminal sequence with 6 histidine residues. The histidine stretch allowed to a rapid and efficient purification of polypeptides derived from the caRLV130 insert sequence on nickel columns (see fig. 6 and table I for denomination and length of 15 polypeptide fragments).

Table I

Definition of nickel column purified CAHSP70 polypeptides. Peptides are coded by pDS56/RBSI-E'-6his recombinant plasmids wherein CARLV130 fragments were cloned. The position refers to nucleotide and aminoacid sequences as shown in Fig. 1. The peptide length refers to the fusion product coded by the recombinant plasmid.

denomination	coding DNA	sequence location on cDNA	position (nt)	length (bp)	denomination	peptide fragment location	position (aa)	length
carLV130/1	whole coding	1-2229	2229	CAHSP70	whole protein	1-656	71-3	664
carLV130/2	3' end cDNA	1393-2229	837	CAHSP70/2	C-terminus	465-656	21-0	202
carLV130/3	5' end cDNA	1-732	732	CAHSP70/3	N-terminus	1-244	28-4	261
carLV130/4	5' end cDNA	1-1027	1027	CAHSP70/4	N-terminus	1-342	39-4	358

After purification, recombinant peptides were used as immunogens to produce mouse immune sera and are therefore able also to induce monoclonal antibodies. Therefore, according the immunization schedule shown in 5 table II, polypeptides, and the whole purified protein as well, induce specific antibodies in a 18-22 g weight Balb/c mouse.

Table II

10 Immunization schedule of 18-22 g weight Balb/c mice with CAHSP70 peptides purified as described in the text and in Fig.6.

Immunogen	Immunization (day 1)	First boost (day 21)	Second boost (day 41)	Serum titer (day 51)
CAHSP70	5 µg	5 µg	10 µg	> 12.800
CAHSP70/2	5 µg	5 µg	10 µg	> 12.800
CAHSP70/3	5 µg	5 µg	10 µg	> 12.800
CAHSP70/4	5 µg	5 µg	10 µg	> 12.800

15 The indicated immunogen concentration was inoculated intraperitoneally in a 200 µl volume. The titer was determined by indirect ELISA with the antigen used for coating at a 200 ng/well concentration, in a final volume of 100 µl, and represents the highest serum dilution able to give an ELISA positive reaction 20 (optical density at 405 nm ≥ two fold the no antigen control value).

Serum titers for each antigen resulted to be > 25 12.800 by immunoenzyme test (indirect ELISA) with the adsorbed antigen at 200 ng/well, in a final volume of 100 µl. The specificity of immunoenzyme test results were confirmed in immunoblot experiments on nitrocellulose filters, as shown in Fig. 7.

The same polypeptides were utilized as immunogens in 30 proliferation assays on peripheral human blood lymphocytes by evaluating the <sup>3</sup>H-thymidine uptake further to 7 day culturing according to standard techniques (7).

Results obtained with different donors (two examples are shown in table III) demonstrate that CAHSP70 is able to induce a good thymidine uptake and the proliferation of naive lymphocytes from umbilical cord blood (Table IV),  
 5 suggesting that the protein itself or parts thereof has a mitogenic activity.

Table III  
 Peripheral blood lymphocytes proliferation induction activity of CAHSP70 and fragments thereof

inducing materials	dose	lymphoproliferative activity <sup>3</sup> H-thymidine uptake (cpm ± SD/2x10 <sup>4</sup> cells)
none	-	500 ± 200
MP-F2	50 µg/ml	13.393 ± 11.555
IL-2	100 U/ml	28.205 ± 18.014
CAHSP70	1 µg/ml	8.730 ± 5.181
CAHSP70/2	1 µg/ml	2.900 ± 2.300
CAHSP70/3	1 µg/ml	3.600 ± 2.700
CAHSP70/4	1 µg/ml	11.685 ± 8.174

Lymphoproliferation of wealthy donor peripheral blood lymphomonocyte cultures further to induction with the CAHSP70 cloned fragments. Positive controls: C. albicans mannoproteic antigen (MP-F2) and Interleukin-2 (IL-2).  
 15 Negative controls: no materials. Shown values represent average values ± SD from 7 experiments with 5 different donors. <sup>3</sup>H-thymidine uptake was determined after 7 days of culture. For technical details, see ref. 7.

Table IV

Umbilical cord blood cell proliferation induction activity of CAHSP70 and fragments thereof

inducing materials	dose	lymphoproliferative activity <sup>3</sup> H-thymidine uptake (cpm ± SD/2x10 <sup>4</sup> cells)	cord blood 1	cord blood 2
none			2.5 ± 0.4	1.3 ± 0.3
IL-2	100 U/ml		37.7 ± 4.5	32.8 ± 6.0
MP-F2	50 µg/ml		3.0 ± 1.4	1.5 ± 0.4
CAHSP70	1 µg/ml		12.5 ± 1.8	22.8 ± 6.6
CAHSP70/2	1 µg/ml		18.2 ± 3.0	23.1 ± 3.9
CAHSP70/3	1 µg/ml		23.8 ± 5.4	20.6 ± 9.2
CAHSP70/4	1 µg/ml		14.8 ± 3.9	17.2 ± 1.7

5

Proliferation of two donor umbilical cord blood cultures further to induction with the CAHSP70 cloned fragments. Positive controls: C. albicans mannoproteic antigen (MP-F2) and Interleukin-2 (IL-2). Negative controls: no materials. Shown values represent average values ± SD from 3 wells. For technical details, see table III legend and ref. 7.

Furthermore, immunoblotting experiments revealed the presence of anti-CAHSP70 antibodies in sera from adult healthy humans, and in particular of the anti-CAHSP70/4 fragment (Fig. 8), suggesting that this fragment contains the immunodominant epitope. Taken together, lymphoproliferations human serum immunoblotting data suggest unequivocally that CAHSP70 is recognized by the immune system during the Candida usual colonization of healthy subjects.

Moreover, in immunoblotting on nitrocellulose filters, anti-CAHSP70 murine sera recognize more than one component of the HSP70 family from heat induced C. albicans extracted proteins (Fig. 5), thus showing that the expression product of caRLV130 insert is a C.

albicans protein which is expressed after the heat shock. According to the above results we named as CAHSP70 the C. albicans protein having the following properties: I) it comprises the aminoacid sequence coded by the caRLV130 insert; II) its gene maps on C. albicans chromosome 1 (having the highest molecular weight); III) its expression is induced by temperature shift; IV) it induces specific antibodies able to recognize cloned and purified fragments (subunits); V) it induces a lymphoproliferation in lymphomonocytic cultures from peripheral human blood. The relevant gene was named as cahsp70.

The CAHSP70 cloning, and its molecular and biochemical characterization, allows to develop a diagnostic molecular method based upon the amplification of DNA inserts corresponding to caRLV130, other than immunological studies of C. albicans 70 kDa heat shock protein expression. According to the caRLV130 insert sequence, we have synthesized oligonucleotides which were utilized for polymerase chain reaction (PCR) experiments, to analyze their ability to amplify DNA fragments which are homologous to C. albicans caRLV130 DNA. Two oligonucleotides (CA2-CA4) were chosen in the regions showing the minimal homology between the caRLV130 cDNA sequence and known HSP70 coding gene sequences (see Fig. 2 for the caRLV130 and YSCSSA1 sequence aligning, see Table V for the definition of minimal homology regions and Table VI for the sequence of oligonucleotides which were utilized for the assay).

The combination of CA2 (GAAATGAAAGATAAGATTGGTGCA) and CA3 (CCACAGTAAATTACCTATTCTTCCTC) oligonucleotides is able to amplify DNA fragments having the expected size and a sequence specific of C. albicans DNA (Fig. 9A), whereas the assay sensitivity is shown in Fig. 9B by

using CA1 (ATGTCTAAAGCTGTTGGTATTG) and CA4 (CTGCACCAATCTTATCTTCATTCACCATCATT) oligonucleotides.

Bibliographic references

- 5 1.Odds. In "Candida and Candidiosis", Bailliere-Tindall,  
London (1988).
- 2.Quinti et al. Clin. Exp. Immunol. 85, 485 (1991).
- 3.Torosantucci et al. J. Infect. Dis. 168, 427 (1993).
- 4.Kaufmann et al. in "The biology of heat shock and  
10 molecular chaperones" R.I. Morimoto et al. Eds. CSHL  
press (1994).
- 5.Morimoto et al. in "The biology of heat shock and  
molecular chaperones" R.I. Morimoto et al. Eds. CSHL  
press (1994).
- 15 6.Stuber et al. Eur. J. Immunol. 220, 819 (1990).
- 7.Ausiello et al. Infect. Immun. 61, 4105 (1993).
- 8.Lasker B.A. et al. Expt. Micol. 16, 155 (1992).

## SEQUENCE LISTING

## (1) GENERAL INFORMATION:

## 5 (i) APPLICANT:

- (A) NAME: Istituto Superiore di Sanita'
- (B) STREET: Viale Regina Elena 299
- (C) CITY: Rome
- (E) COUNTRY: Italy

10 (F) POSTAL CODE (ZIP): 00161

(A) NAME: Universita' degli Studi di Roma La Sapienza

(B) STREET: P.le Aldo Moro 5

(C) CITY: Rome

15 (E) COUNTRY: Italy

(F) POSTAL CODE (ZIP): 00184

(ii) TITLE OF INVENTION: Candida heath shock protein, gene and  
uses thereof

20 (iii) NUMBER OF SEQUENCES: 2

(iv) COMPUTER READABLE FORM:

- (A) MEDIUM TYPE: Floppy disk
- (B) COMPUTER: IBM PC compatible
- (C) OPERATING SYSTEM: PC-DOS/MS-DOS
- (D) SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPO)

## (2) INFORMATION FOR SEQ ID NO: 1:

30 (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 2001 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

## 35 (ix) FEATURE:

- (A) NAME/KEY: CDS
- (B) LOCATION: 1..1968

40

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1:

	ATG TCT AAA GCT GTT GGT ATT GAT TTA GGT ACA ACC TAT TCT TGT GTT	48
	Met Ser Lys Ala Val Gly Ile Asp Leu Gly Thr Thr Tyr Ser Cys Val	
5	1                   5                   10                   15	
	GCT CAT TTT GCC AAT GAT AGA GTT GAA ATT ATT GCT AAT GAT CAA GGT	96
	Ala His Phe Ala Asn Asp Arg Val Glu Ile Ile Ala Asn Asp Gln Gly	
	20                   25                   30	
10	AAT AGA ACT ACC CCT TCA TTT GTT GCC TTC ACT GAT ACT GAA AGA TTG	144
	Asn Arg Thr Thr Pro Ser Phe Val Ala Phe Thr Asp Thr Glu Arg Leu	
	35                   40                   45	
15	ATT GGT GAT GCT GCC AAG AAT CAA GCT GCT ATG AAC CCA GCA AAC ACT	192
	Ile Gly Asp Ala Ala Lys Asn Gln Ala Ala Met Asn Pro Ala Asn Thr	
	50                   55                   60	
20	GTT TTC GAT GCT AAA CGT TTA ATT GGG AGA AAA TTT GAT GAT CCA GAA	240
	Val Phe Asp Ala Lys Arg Leu Ile Gly Arg Lys Phe Asp Asp Pro Glu	
	65                   70                   75                   80	
25	GTT ATA AAT GAT GCT AAA CAT TTC CCA TTT AAA GTC ATT GAT AAA GCA	288
	Val Ile Asn Asp Ala Lys His Phe Pro Phe Lys Val Ile Asp Lys Ala	
	85                   90                   95	
30	GGT AAA CCA GTG ATT CAA GTT GAA TAT AAA GGT GAA ACT AAA ACT TTT	336
	Gly Lys Pro Val Ile Gln Val Glu Tyr Lys Gly Glu Thr Lys Thr Phe	
	100                  105                  110	
35	TCA CCA GAA GAA ATT TCT TCA ATG GTT TTA ACA AAA ATG AAA GAA ATT	384
	Ser Pro Glu Glu Ile Ser Ser Met Val Leu Thr Lys Met Lys Glu Ile	
	115                  120                  125	
40	GCT GAA GGT TAT TTG GGT TCT ACT GTT AAA GAT GCT GTT GTT ACT GTT	432
	Ala Glu Gly Tyr Leu Gly Ser Thr Val Lys Asp Ala Val Val Thr Val	
	130                  135                  140	
45	CCA GCT TAT TTC AAT GAT TCT CAA AGA CAA GCC ACC AAA GAT GCT GGT	480
	Pro Ala Tyr Phe Asn Asp Ser Gln Arg Gln Ala Thr Lys Asp Ala Gly	
	145                  150                  155                  160	
	ACT ATT GCT GGT TTG AAT GTT TTA AGA ATT ATT AAT GAA CCT ACT GCT	528
	Thr Ile Ala Gly Leu Asn Val Leu Arg Ile Ile Asn Glu Pro Thr Ala	
	165                  170                  175	

	GCT GCC ATT GCT TAT GGT TTA GAT AAA AAA GGT TCC AGA GGT GAA CAT	576
	Ala Ala Ile Ala Tyr Gly Leu Asp Lys Lys Gly Ser Arg Gly Glu His	
180	185	190
5		
	AAT GTT TTA ATT TTC GAT TTG GGT GGT ACT TTT GAT GTT TCA TTA	624
	Asn Val Leu Ile Phe Asp Leu Gly Gly Thr Phe Asp Val Ser Leu	
195	200	205
10		
	TTA GCC ATT GAT GAA GGT ATT TTC GAA GTT AAA GCC ACT GCT GGT GAT	672
	Leu Ala Ile Asp Glu Gly Ile Phe Glu Val Lys Ala Thr Ala Gly Asp	
210	215	220
15		
	ACT CAT TTG GGT GGT GAA GAT TTT GAT AAC AGA TTA GTC AAC TTC TTT	720
	Thr His Leu Gly Gly Glu Asp Phe Asp Asn Arg Leu Val Asn Phe Phe	
225	230	235
20		
	ATT CAA GAA TTC AAG AGA AAG AAC AAG AAA GAT ATT TCC ACC AAC CAA	768
	Ile Gln Glu Phe Lys Arg Lys Asn Lys Lys Asp Ile Ser Thr Asn Gln	
245	250	255
25		
	AGA GCT TTA AGA AGA TTA AGA ACT GCT TGT GAA AGA GCC AAG AGA ACT	816
	Arg Ala Leu Arg Arg Leu Arg Thr Ala Cys Glu Arg Ala Lys Arg Thr	
260	265	270
30		
	TTG TCT TCT GCT CAA ACC TCA ATT GAA ATT GAT TCC TTA TAT GAA	864
	Leu Ser Ser Ser Ala Gln Thr Ser Ile Glu Ile Asp Ser Leu Tyr Glu	
275	280	285
35		
	GGT ATT GAC TTC TAC ACT TCA ATC ACC AGA GCC AGA TTT GAA GAA TTG	912
	Gly Ile Asp Phe Tyr Thr Ser Ile Thr Arg Ala Arg Phe Glu Glu Leu	
290	295	300
40		
	TGT GCT GAC TTG TTT AGA TCC ACT TTA GAT CCA GTT GGT AAA GTT TTA	960
	Cys Ala Asp Leu Phe Arg Ser Thr Leu Asp Pro Val Gly Lys Val Leu	
305	310	315
45		
	GCT GAT GCC AAG ATT GAT AAA TCT CAA GTT GAA GAA ATT GTC TTG GTT	1008
	Ala Asp Ala Lys Ile Asp Lys Ser Gln Val Glu Glu Ile Val Leu Val	
325	330	335
	GGT GGG TCC ACC AGA ATT CCA AAG ATT CAA AAA TTG GTT TCT GAT TTC	1056
	Gly Gly Ser Thr Arg Ile Pro Lys Ile Gln Lys Leu Val Ser Asp Phe	
340	345	350

	TTT AAT GGT AAA GAA TTG AAT AAA TCT ATC AAC CCT GAT GAA GCT GTT Phe Asn Gly Lys Leu Asn Lys Ser Ile Asn Pro Asp Glu Ala Val	355	360	365	1104
5	GCT TAT GGT GCT GCT CAA GCT GCC ATT TTA ACT GGT GAT ACT TCT Ala Tyr Gly Ala Ala Val Gln Ala Ala Ile Leu Thr Gly Asp Thr Ser	370	375	380	1152
10	TCC AAG ACT CAA GAT ATT TTG TTA TTG GAT GTT GCT CCA TTG TCA TTA Ser Lys Thr Gln Asp Ile Leu Leu Asp Val Ala Pro Leu Ser Leu	385	390	395	1200
15	GGT ATT GAA ACT GCT GGT GGT ATC ATG ACC AAA TTG ATT CCA AGA AAT Gly Ile Glu Thr Ala Gly Ile Met Thr Lys Leu Ile Pro Arg Asn	405	410	415	1248
20	TCT ACT ATT CCA ACT AAG AAA TCA GAA ACT TTC TCC ACT TAT GCC GAT Ser Thr Ile Pro Thr Lys Ser Glu Thr Phe Ser Thr Tyr Ala Asp	420	425	430	1296
25	AAC CAA CCA GGT GTT TTG ATT CAA GTG TTT GAA GGT GAA AGA GCT AAA Asn Gln Pro Gly Val Leu Ile Gln Val Phe Glu Gly Glu Arg Ala Lys	435	440	445	1344
30	ACT AAA GAT AAC AAC TTG TTG GGT AAA TTT GAA TTA TCT GGT ATT CCA Thr Lys Asp Asn Asn Leu Leu Gly Lys Phe Glu Leu Ser Gly Ile Pro	450	455	460	1392
35	CCA GCT CCA AGA GGC GTC CCT CAA ATT GAA GTT ACT TTC GAT ATT GAT Pro Ala Pro Arg Gly Val Pro Gln Ile Glu Val Thr Phe Asp Ile Asp	465	470	475	1440
40	GCT AAT GGT ATC TTG AAT GTT TCT GCT TTA GAA AAA GGT ACT GGT AAA Ala Asn Gly Ile Leu Asn Val Ser Ala Leu Glu Lys Gly Thr Gly Lys	485	490	495	1488
45	ACT CAA AAG ATT ACT ATC ACC AAC GAT AAA GGT AGA TTA TCC AAA GAA Thr Gln Lys Ile Thr Ile Thr Asn Asp Lys Gly Arg Leu Ser Lys Glu	500	505	510	1536
50	GAA ATT GAT AAA ATG GTT AGT GAA GCT GAA AAA TTC AAA GAA GAA GAT Glu Ile Asp Lys Met Val Ser Glu Ala Glu Lys Phe Lys Glu Glu Asp	515	520	525	1584

	GAA AAG GAA GCT GCT AGA GTC CAA GCC AAG AAT CAA TTG GAA TCT TAT	1632
	Glu Lys Glu Ala Ala Arg Val Gln Ala Lys Asn Gln Leu Glu Ser Tyr	
	530 535 540	
5	GCT TAT TCA TTG AAA AAC ACA ATC AAT GAT GGT GAA ATG AAA GAT AAG	1680
	Ala Tyr Ser Leu Lys Asn Thr Ile Asn Asp Gly Glu Met Lys Asp Lys	
	545 550 555 560	
10	ATT GGT GCA GAT GAT AAA GAA AAA TTA ACT AAA GCC ATT GAT GAA ACT	1728
	Ile Gly Ala Asp Asp Lys Glu Lys Leu Thr Lys Ala Ile Asp Glu Thr	
	565 570 575	
15	ATT TCT TGG TTA GAT GCA TCT CAA GCT GCT TCT ACT GAA GAA TAC GAA	1776
	Ile Ser Trp Leu Asp Ala Ser Gln Ala Ala Ser Thr Glu Glu Tyr Glu	
	580 585 590	
	GAT AAA CGT AAA GAA TTA GAA TCA GTT GCT AAT CCA ATC ATT AGT GGT	1824
	Asp Lys Arg Lys Glu Leu Glu Ser Val Ala Asn Pro Ile Ile Ser Gly	
	595 600 605	
20	GCT TAT GGT GCT GCC GGT GGC GCT CCA GGT GGT GCA GGC GGA TTC CCA	1872
	Ala Tyr Gly Ala Ala Gly Gly Ala Pro Gly Gly Ala Gly Phe Pro	
	610 615 620	
25	GGT GCT GGT GGC TTC CCA GGT GCC CCA GGT GCC GGT GGT CCA GGT	1920
	Gly Ala Gly Gly Phe Pro Gly Gly Ala Pro Gly Ala Gly Gly Pro Gly	
	625 630 635 640	
30	GGT GCT ACT GGT GGT GAA TCT AGT GGA CCA ACT GTT GAA GAA GTT GAT	1968
	Gly Ala Thr Gly Gly Glu Ser Ser Gly Pro Thr Val Glu Glu Val Asp	
	645 650 655	
35	TAAATGAGGAAGAAATAGGTAAATTACTGTGG	2000

## (2) INFORMATION FOR SEQ ID NO: 2:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 656 amino acids
  - (B) TYPE: amino acid
  - (D) TOPOLOGY: linear

- (ii) MOLECULE TYPE: protein

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2:

Met Ser Lys Ala Val Gly Ile Asp Leu Gly Thr Thr Tyr Ser Cys Val  
1 5 10 15

5 Ala His Phe Ala Asn Asp Arg Val Glu Ile Ile Ala Asn Asp Gln Gly  
20 25 30

Asn Arg Thr Thr Pro Ser Phe Val Ala Phe Thr Asp Thr Glu Arg Leu  
35 40 45

Ile Gly Asp Ala Ala Lys Asn Gln Ala Ala Met Asn Pro Ala Asn Thr  
10 50 55 60

Val Phe Asp Ala Lys Arg Leu Ile Gly Arg Lys Phe Asp Asp Pro Glu  
65 70 75 80

Val Ile Asn Asp Ala Lys His Phe Pro Phe Lys Val Ile Asp Lys Ala  
85 90 95

15 Gly Lys Pro Val Ile Gln Val Glu Tyr Lys Gly Glu Thr Lys Thr Phe  
100 105 110

Ser Pro Glu Glu Ile Ser Ser Met Val Leu Thr Lys Met Lys Glu Ile  
115 120 125

Ala Glu Gly Tyr Leu Gly Ser Thr Val Lys Asp Ala Val Val Thr Val  
20 130 135 140

Pro Ala Tyr Phe Asn Asp Ser Gln Arg Gln Ala Thr Lys Asp Ala Gly  
145 150 155 160

Thr Ile Ala Gly Leu Asn Val Leu Arg Ile Ile Asn Glu Pro Thr Ala  
165 170 175

25 Ala Ala Ile Ala Tyr Gly Leu Asp Lys Lys Gly Ser Arg Gly Glu His  
180 185 190

Asn Val Leu Ile Phe Asp Leu Gly Gly Thr Phe Asp Val Ser Leu  
195 200 205

Leu Ala Ile Asp Glu Gly Ile Phe Glu Val Lys Ala Thr Ala Gly Asp  
30 210 215 220

Thr His Leu Gly Gly Glu Asp Phe Asp Asn Arg Leu Val Asn Phe Phe  
225 230 235 240

Ile Gln Glu Phe Lys Arg Lys Asn Lys Lys Asp Ile Ser Thr Asn Gln  
245 250 255

35 Arg Ala Leu Arg Arg Leu Arg Thr Ala Cys Glu Arg Ala Lys Arg Thr  
260 265 270

Leu Ser Ser Ser Ala Gln Thr Ser Ile Glu Ile Asp Ser Leu Tyr Glu  
275 280 285

Gly Ile Asp Phe Tyr Thr Ser Ile Thr Arg Ala Arg Phe Glu Leu  
40

	290	295	300
	Cys Ala Asp Leu Phe Arg Ser Thr Leu Asp Pro Val Gly Lys Val Leu		
	305	310	315
	Ala Asp Ala Lys Ile Asp Lys Ser Gln Val Glu Glu Ile Val Leu Val		320
5		325	330
	Gly Gly Ser Thr Arg Ile Pro Lys Ile Gln Lys Leu Val Ser Asp Phe		
	340	345	350
	Phe Asn Gly Lys Glu Leu Asn Lys Ser Ile Asn Pro Asp Glu Ala Val		
	355	360	365
10	Ala Tyr Gly Ala Ala Val Gln Ala Ala Ile Leu Thr Gly Asp Thr Ser		
	370	375	380
	Ser Lys Thr Gln Asp Ile Leu Leu Asp Val Ala Pro Leu Ser Leu		
	385	390	395
	Gly Ile Glu Thr Ala Gly Gly Ile Met Thr Lys Leu Ile Pro Arg Asn		400
15		405	410
	Ser Thr Ile Pro Thr Lys Ser Glu Thr Phe Ser Thr Tyr Ala Asp		
	420	425	430
	Asn Gln Pro Gly Val Leu Ile Gln Val Phe Glu Gly Glu Arg Ala Lys		
	435	440	445
20	Thr Lys Asp Asn Asn Leu Leu Gly Lys Phe Glu Leu Ser Gly Ile Pro		
	450	455	460
	Pro Ala Pro Arg Gly Val Pro Gln Ile Glu Val Thr Phe Asp Ile Asp		
	465	470	475
	Ala Asn Gly Ile Leu Asn Val Ser Ala Leu Glu Lys Gly Thr Gly Lys		480
25		485	490
	Thr Gln Lys Ile Thr Ile Thr Asn Asp Lys Gly Arg Leu Ser Lys Glu		
	500	505	510
	Glu Ile Asp Lys Met Val Ser Glu Ala Glu Lys Phe Lys Glu Glu Asp		
	515	520	525
30	Glu Lys Glu Ala Ala Arg Val Gln Ala Lys Asn Gln Leu Glu Ser Tyr		
	530	535	540
	Ala Tyr Ser Leu Lys Asn Thr Ile Asn Asp Gly Glu Met Lys Asp Lys		
	545	550	555
	Ile Gly Ala Asp Asp Lys Glu Lys Leu Thr Lys Ala Ile Asp Glu Thr		560
35		565	570
	Ile Ser Trp Leu Asp Ala Ser Gln Ala Ala Ser Thr Glu Glu Tyr Glu		
	580	585	590
	Asp Lys Arg Lys Glu Leu Glu Ser Val Ala Asn Pro Ile Ile Ser Gly		
	595	600	605
40	Ala Tyr Gly Ala Ala Gly Gly Ala Pro Gly Gly Ala Gly Phe Pro		

610                    615                    620  
Gly Ala Gly Gly Phe Pro Gly Gly Ala Pro Gly Ala Gly Gly Pro Gly  
625                    630                    635                    640  
Gly Ala Thr Gly Gly Glu Ser Ser Gly Pro Thr Val Glu Glu Val Asp  
5                      645                    650                    655

## Claims

1. A nucleic acid comprising a nucleotide sequence coding the protein having the amino acid sequence of SEQ ID No.2 or parts thereof.

2. A nucleic acid comprising a nucleotide sequence with at least a 65% homology with the nucleotide sequence of SEQ ID No.1 or parts thereof.

3. A nucleic acid according to claim 2 comprising the nucleotide sequence of SEQ ID No.1 or parts thereof.

4. Composition comprising a nucleic acid according to any of claims 1 to 3.

5. Use of the nucleic acid according to any of claims 1 to 3 for oligonucleotide probes to be used in diagnosis and typing of Candida and Candida related pathologies.

6. Oligonucleotide having a sequence comprised in SEQ ID No. 1 to be used for PCR (polymerase chain reaction) to detect the presence in biological and/or environment samples either of C. albicans or of other Candida species or of yeast-like related microorganisms comprising said gene and/or in a labeled form (radionuclides, biotin, enzymes, etc.) to detect the presence in biological and/or environment samples either of C. albicans or of other related and/or for the C. albicans or related species typing and/or diagnosis and/or as potential antibiotic and/or chemiotherapeutic targets, or antisense RNA active for Candida species and/or yeast-like related microorganisms coding an homologous sequence.

7. Polypeptide having the aminoacid sequence comprised in the SEQ ID No.1, or having at least a 50% homology with SEQ ID No. 1 or fragments, and/or functional and immunologic homologous thereof.

atg tct aaa gct gtt ggt att gat tta cgt acc tat ttc tgt gtc cat ttt gcc eet gat aga gtt ctt gaa att att gct eet gat caa ggt aat aga act acc cct  
 H S K A V G I D L G T Y S C V A H F A N D R V E I I A N D Q G N R T T P>  
 S F V N F T D T E R L I C D A A K N P A M N P V F D A K R L I G R>  
 100  
 tca ltt gtt gcc ttc act gnt act gat gaa ttt ggt gat gct gcc eag aat caa gct gat atg aac cca gca aac act gtt ttc gat gct gat gaa cgt tta att ggg aga  
 K F H N F V I N D A K II F P F K V I D K A V P V T Q V F Y K G E T K T>  
 200  
 aaa tct gat gat cca gaa gtt aat gat gct gaa cat ttc cca ltt aaa gtc att gat gaa gca ggt aat cca gtt gaa tat aat ggt gaa act aaa act  
 F S P E F J S H V L T K H K E I A E G Y L G S T V K D A V V P A Y F>  
 300  
 ttc tca cca gaa gaa att lct tca atg gtt tta aca aaa gaa att gct gaa ggt tat ttc ggt tct act gtt gtt act gtt cca gct tat ttc  
 F S P E F J S H V L T K H K E I A E G Y L G S T V K D A V V P A Y F>  
 400  
 aat gat tct can aga caa gcc acc aaa gat gct ggt act att gct ggt ttc aat gtt tta aga att att aat gaa cct act gct gtt gcc att gct tat ggt tta gal aaa  
 N D S Q R O A T K D A G T I A G L N V L R I I N E P T A A Y G L D K>  
 500  
 aaa ggt tcc aga ggt gaa cat aat gtt tta att ttc gat ttc ggt ggt act ttc gat gtt tca tta ttc gcc att gat gaa ggt att ttc gaa gtt aaa gcc act gct  
 K C S R G E H N V L I F D L G G T F D V S L A I D E G I F E V K A T A>  
 600  
 ggt gat act cat tlg ggt ggt gaa gat ttc gat aac aga ttc gtc eac ttc ttc att cca gaa ttc wgg aga aag aac gat att tcc acc aac caa ega gct tca  
 G D T H L G C E D F N R L V N F I Q E F K N K D I S T N Q R A L>  
 700  
 ngt gat act cat tlg ggt ggt gaa gat ttc gat aac aga ttc gtc eac ttc ttc att cca gaa att gat tcc tta tat gaa ggt att gac ttc tac act tca atc  
 G R R I H T A C E R A K R T L S S A Q T S I D S L Y E G I D F Y T S I>  
 800  
 acc aga tta ngt act gct tgt gaa gaa ttc gat gtc act tcc act ttc tct tct gct caa acc tca att gaa att gat tcc tta tat gaa ggt att gat oaa tct caa gtt gaa gaa att  
 T R A R F E E L C A D L F R S T L D P V G K V L A D K S Q V E E I>  
 900

1 / 12

**SUBSTITUTE SHEET (RULE 26)**

**SUBSTITUTE SHEET (RULE 26)**

FIG. 2

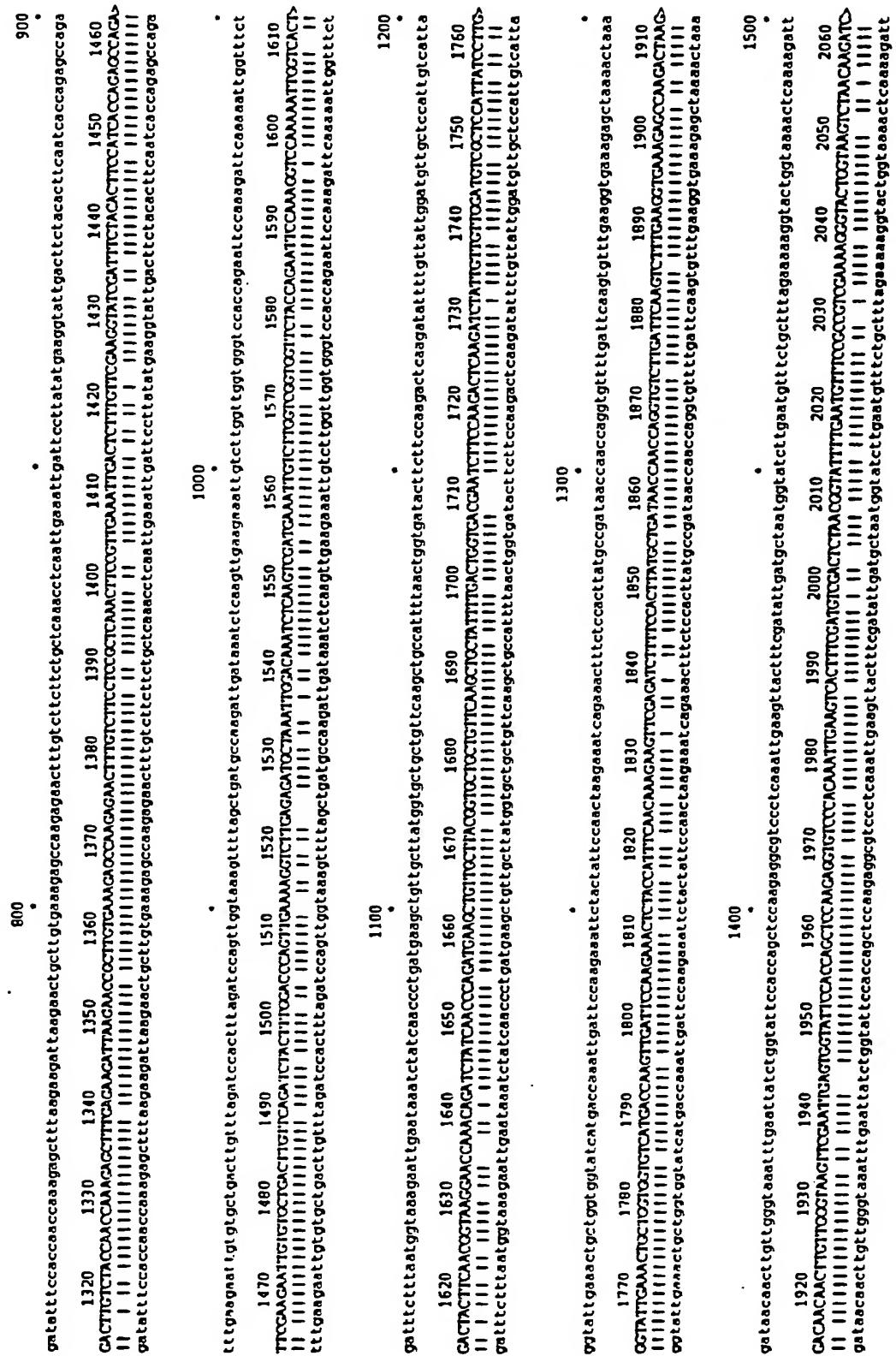


FIG. 2 (cont.)

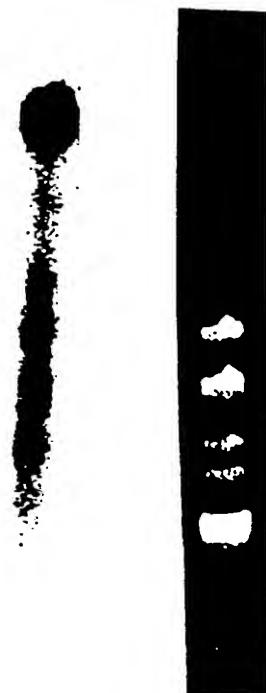
FIG. 2 (cont.)

The graph displays the distribution of sequence lengths across different sequence types. The x-axis represents the sequence length, ranging from 100 to 600. The y-axis represents the count of sequences, ranging from 0 to 400. The data points are as follows:

Sequence Type	Length Range (approx.)	Count (approx.)
makavg	100-150	100
makavg	150-200	20
makavg	200-250	30
makavg	250-300	40
makavg	300-350	50
makavg	350-400	60
makavg	400-450	70
makavg	450-500	80
makavg	500-550	90
makavg	550-600	100
sqrqatkdegtlaginvrlneptaa	100-150	160
sqrqatkdegtlaginvrlneptaa	150-200	170
sqrqatkdegtlaginvrlneptaa	200-250	190
sqrqatkdegtlaginvrlneptaa	250-300	200
sqrqatkdegtlaginvrlneptaa	300-350	210
sqrqatkdegtlaginvrlneptaa	350-400	220
sqrqatkdegtlaginvrlneptaa	400-450	230
sqrqatkdegtlaginvrlneptaa	450-500	240
sqrqatkdegtlaginvrlneptaa	500-550	250
sqrqatkdegtlaginvrlneptaa	550-600	260
sqrqatkdegtlaginvrlneptaa	600-650	270
sqrqatkdegtlaginvrlneptaa	650-700	280
sqrqatkdegtlaginvrlneptaa	700-750	290
feelcad	100-150	300
feelcad	150-200	310
feelcad	200-250	320
feelcad	250-300	330
feelcad	300-350	340
feelcad	350-400	350
feelcad	400-450	360
feelcad	450-500	370
feelcad	500-550	380
feelcad	550-600	390
feelcad	600-650	400
feelcad	650-700	410
feelcad	700-750	420
feelcad	750-800	430
feelcad	800-850	440
dnlugkfelegtrpapgvulevfdldangl	100-150	450
dnlugkfelegtrpapgvulevfdldangl	150-200	460
dnlugkfelegtrpapgvulevfdldangl	200-250	470
dnlugkfelegtrpapgvulevfdldangl	250-300	480
dnlugkfelegtrpapgvulevfdldangl	300-350	490
dnlugkfelegtrpapgvulevfdldangl	350-400	500
dnlugkfelegtrpapgvulevfdldangl	400-450	510
dnlugkfelegtrpapgvulevfdldangl	450-500	520
dnlugkfelegtrpapgvulevfdldangl	500-550	530
dnlugkfelegtrpapgvulevfdldangl	550-600	540
dnlugkfelegtrpapgvulevfdldangl	600-650	550
dnlugkfelegtrpapgvulevfdldangl	650-700	560
dnlugkfelegtrpapgvulevfdldangl	700-750	570
dnlugkfelegtrpapgvulevfdldangl	750-800	580
dnlugkfelegtrpapgvulevfdldangl	800-850	590

三  
G.  
E

A              B



« 3,5 Mbp

« 0,7 Mbp

FIG. 4

7/12

SUBSTITUTE SHEET (RULE 26)

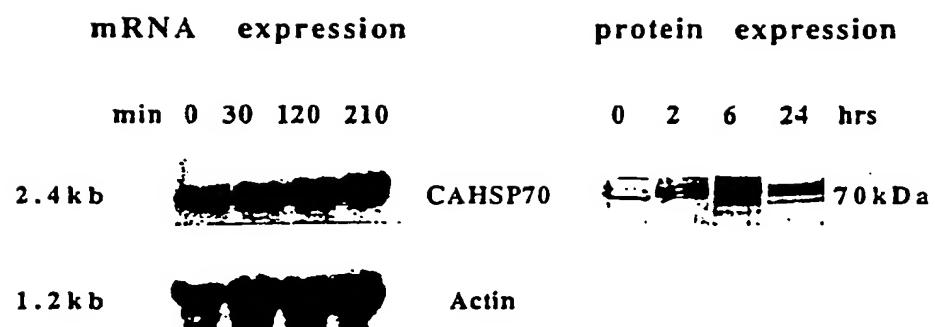
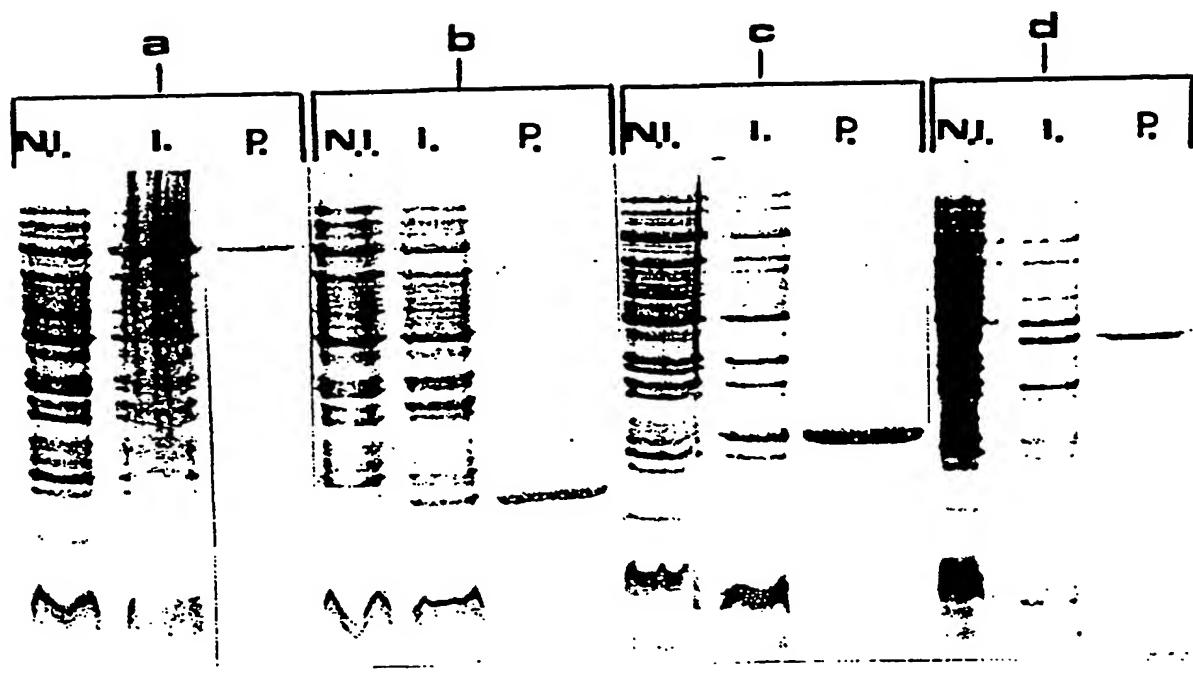


FIG. 5

8/12

**SUBSTITUTE SHEET (RULE 26)**

A



B



CAHSP70	N-6xHis-	<b>a</b>	C	~71kDa
CAHSP70/2		N-6xHis-	<b>b</b>	C ~21kDa
CAHSP70/3	N-6xHis-	<b>c</b>	-C	~28kDa
CAHSP70/4	N-6xHis-	<b>d</b>	-C	~39kDa

FIG. 6

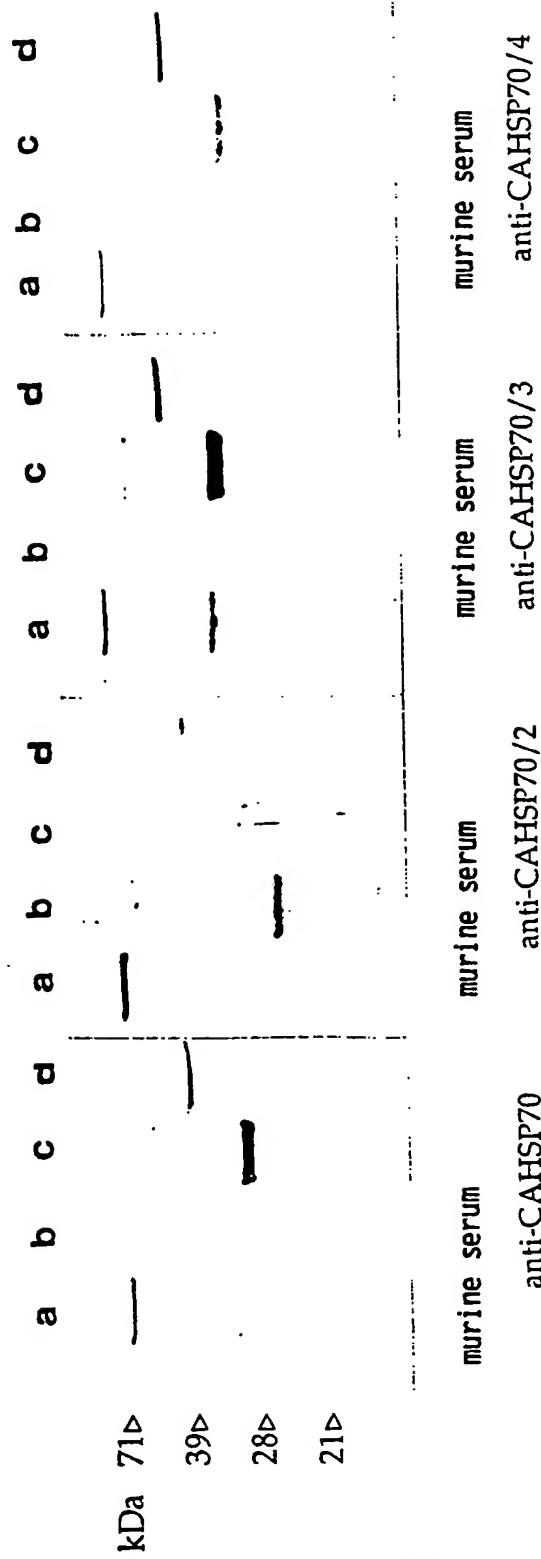


FIG. 7

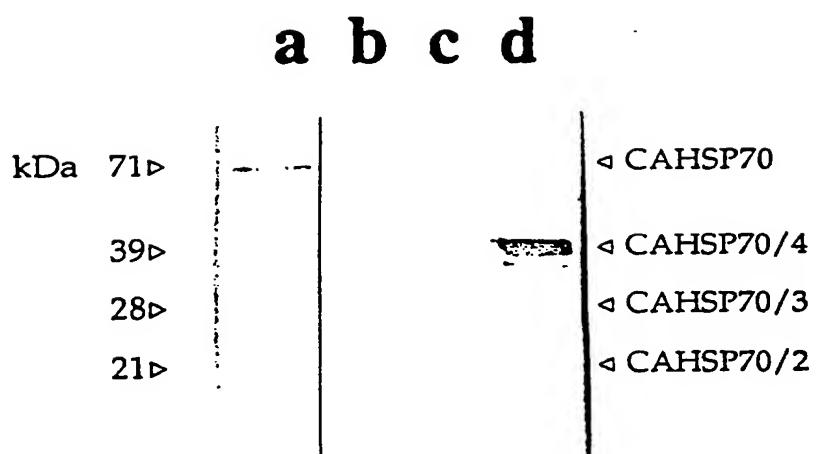
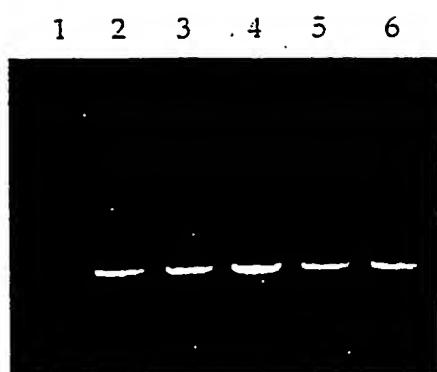


FIG. 8

**A**

△ 335 bp

△ oligonucleotides

**B**

△ 1690 bp

FIG. 9

## INTERNATIONAL SEARCH REPORT

International Application No
PC1, IT 96/00097

A. CLASSIFICATION OF SUBJECT MATTER  
 IPC 6 C12N15/11 C12Q1/68  
 A61K39/00

C07K14/40

C07K16/14

G01N33/569

According to International Patent Classification (IPC) or to both national classification and IPC

## B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)  
 IPC 6 C12N C12Q C07K G01N A61K

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practical, search terms used)

## C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	NUCLEIC ACIDS RESEARCH, vol. 17, no. 2, 25 January 1989, OXFORD GB, pages 805-806, XP002011023 M.R.SLATER AND E.A.CRAIG: "The SSA1 and SSA2 genes of the yeast <i>Saccharomyces</i> <i>cerevisiae</i> " see figure 1 --- A EP,A,0 406 029 (J.P.BURNIE AND R.C.MATTHEWS) 2 January 1991 see the whole document ---	1-4,7,8  9-12 -/-

 Further documents are listed in the continuation of box C. Patent family members are listed in annex.

## \* Special categories of cited documents :

- \*A\* document defining the general state of the art which is not considered to be of particular relevance
- \*E\* earlier document but published on or after the international filing date
- \*L\* document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)
- \*O\* document referring to an oral disclosure, use, exhibition or other means
- \*P\* document published prior to the international filing date but later than the priority date claimed

\*T\* later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

\*X\* document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

\*Y\* document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.

\*&amp;\* document member of the same patent family

2

Date of the actual completion of the international search

19 August 1996

Date of mailing of the international search report

28.08.96

Name and mailing address of the ISA

European Patent Office, P.B. 5818 Patentlaan 2  
 NL - 2280 HV Rijswijk  
 Tel. (+ 31-70) 340-2040, Tx. 31 651 epo nl,  
 Fax (+ 31-70) 340-3016

Authorized officer

Cupido, M

## INTERNATIONAL SEARCH REPORT

Int'l Application No  
PCT/IT 96/00097

## C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

Category	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
P,X	INFECTION AND IMMUNITY, vol. 63 , no. 10, October 1995, WASHINGTON US, pages 4039-4045, XP002011024 R.LA VALLE ET AL.: "Molecular cloning and expression of a 70-kilodalton heat shock protein of <i>Candida albicans</i> " see the whole document ---	1-12
A	JOURNAL OF MOLECULAR EVOLUTION, vol. 38, no. 1, January 1994, pages 1-17, XP000578395 W.R.BORSTEIN ET AL.: "Molecular evolution of the HSP70 multigene family". see table 3 -----	1-4,7,8

FURTHER INFORMATION CONTINUED FROM PCT/ISA/210

Remark : Although claim 12, insofar as it relates to *in vivo* uses, is directed to a method of treatment of the human/animal body the search has been carried out and based on the alleged effects of the composition.

**INTERNATIONAL SEARCH REPORT**

I: International application No.

PCT/IT 96/00097

**Box I Observations where certain claims were found unsearchable (Continuation of item 1 of first sheet)**

This international search report has not been established in respect of certain claims under Article 17(2)(a) for the following reasons:

1.  Claims Nos.: because they relate to subject matter not required to be searched by this Authority, namely:  
Please see Further Information sheet enclosed.
2.  Claims Nos.: because they relate to parts of the international application that do not comply with the prescribed requirements to such an extent that no meaningful international search can be carried out, specifically:
3.  Claims Nos.: because they are dependent claims and are not drafted in accordance with the second and third sentences of Rule 6.4(a).

**Box II Observations where unity of invention is lacking (Continuation of item 2 of first sheet)**

This International Searching Authority found multiple inventions in this international application, as follows:

1.  As all required additional search fees were timely paid by the applicant, this international search report covers all searchable claims.
2.  As all searchable claims could be searched without effort justifying an additional fee, this Authority did not invite payment of any additional fee.
3.  As only some of the required additional search fees were timely paid by the applicant, this international search report covers only those claims for which fees were paid, specifically claims Nos.:
4.  No required additional search fees were timely paid by the applicant. Consequently, this international search report is restricted to the invention first mentioned in the claims; it is covered by claims Nos.:

**Remark on Protest**

The additional search fees were accompanied by the applicant's protest.  
 No protest accompanied the payment of additional search fees.

## INTERNATIONAL SEARCH REPORT

Information on patent family members

International Application No  
PCT/IT 96/00097

Patent document cited in search report	Publication date	Patent family member(s)		Publication date
EP-A-0406029	02-01-91	AT-T-	120490	15-04-95
		AU-B-	640394	26-08-93
		AU-B-	6036290	17-01-91
		CA-A-	2034504	31-12-90
		DE-D-	69018142	04-05-95
		DE-T-	69018142	27-07-95
		ES-T-	2072393	16-07-95
		WO-A-	9100351	10-01-91
		GB-A,B	2240979	21-08-91
		JP-T-	4502257	23-04-92
		US-A-	5541077	30-07-96
		US-A-	5288639	22-02-94